(12) **UK Patent Application** (19) **GB** (11) **2 387 676** (13) **A**

(43) Date of A Publication **22.10.2003**

(21) Application No **0208607.2**

(22) Date of Filing **15.04.2002**

(71) Applicant(s)
Agena Limited
(Incorporated in the United Kingdom)
11 Main Street, Caldecote, Cambridge,
CB3 7NU, United Kingdom

(72) Inventor(s)
Norman Fenton
Martin Neil

(74) Agent and/or Address for Service
Mathys & Squire
100 Grays Inn Road, LONDON, WC1X 8AL,
United Kingdom

(51) INT CL⁷
G06F 17/60

(52) UK CL (Edition V)
G4A AUXP AUXX

(56) Documents Cited
EP   1192802 A   EP   1155571 A
WO 2002/011445 A   WO 2002/006757 A
WO 2002/004295 A   US   5704017 A

(58) Field of Search
UK CL (Edition V) **G4A** AUXP AUXX
INT CL⁷ **G06F** 17/00 17/60 19/00, **H04N**
Other: **ON-LINE:EPODOC, WPI, JAPIO**

(54) Abstract Title
A method of recommending television programmes

(57) A system for recommending television programmes has a programme classifier deriving membership functions indicating the degree of membership each programme has of the classes for each programme attribute; a viewer profiler monitoring which programmes are watched by the individual viewer and learning through a Bayesian network a preference profile for the individual as a function of those classes; and a programme recommender serving to recommend to the viewer those available programmes whose membership functions most closely match the preference profile.
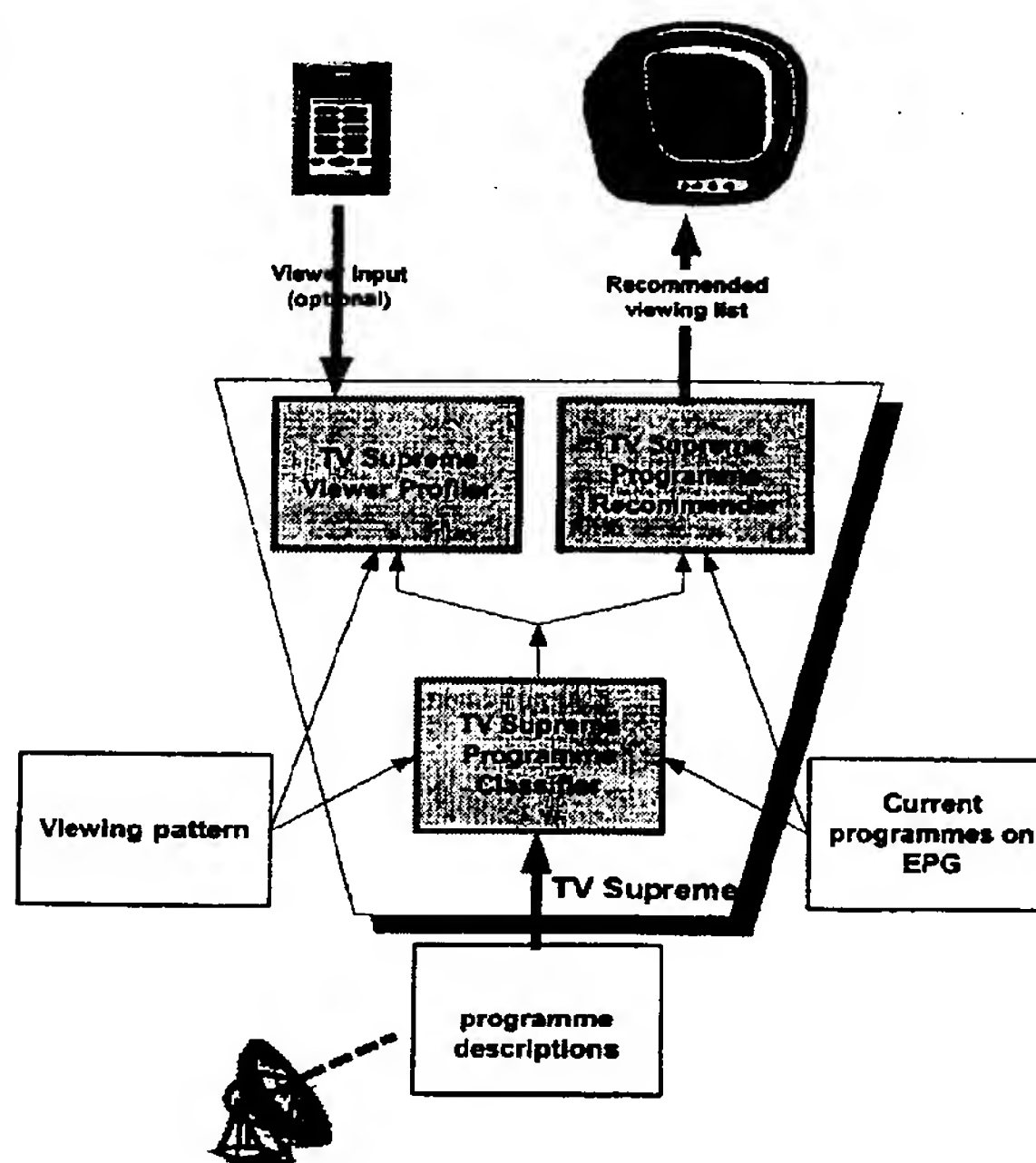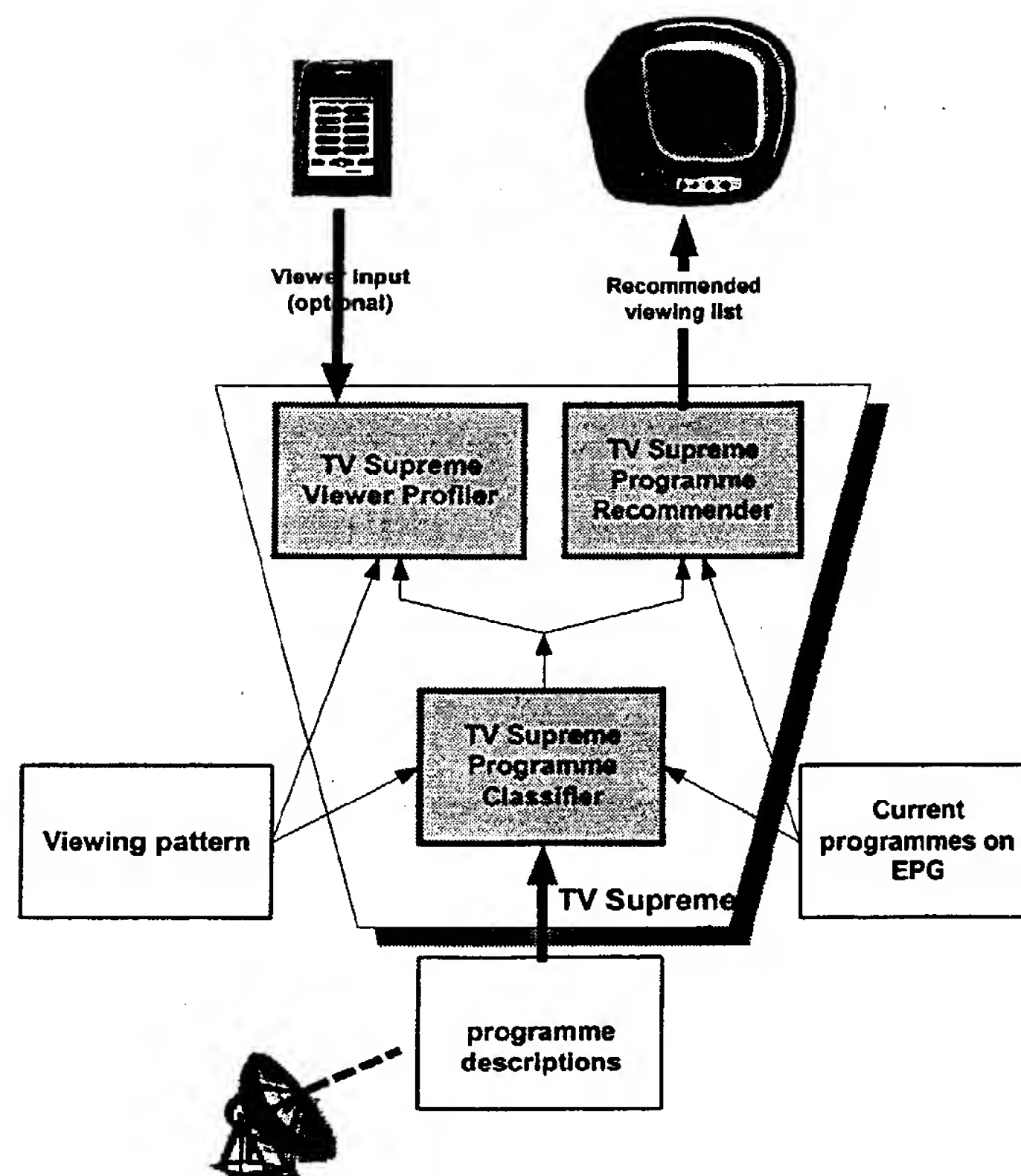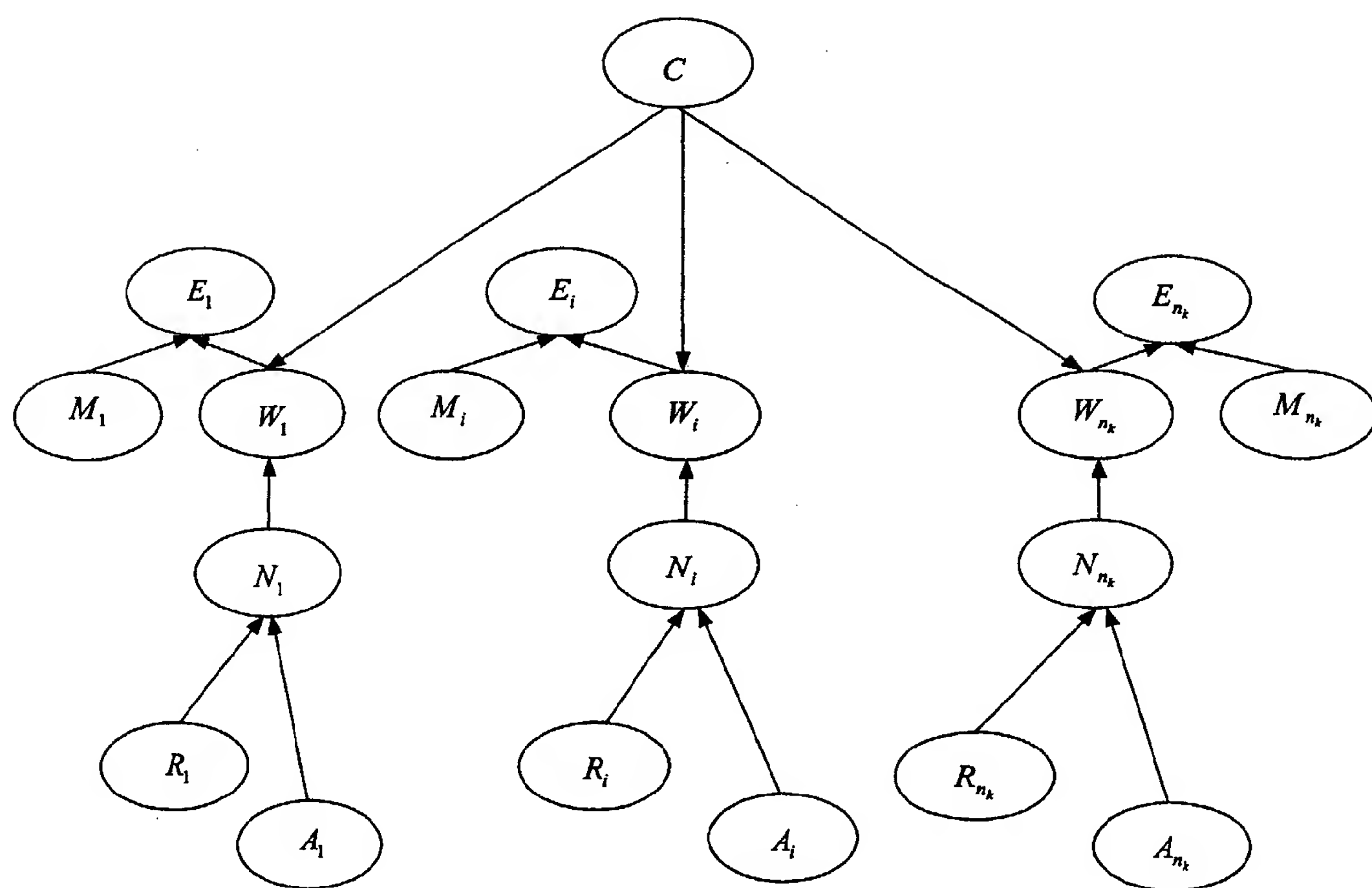
Fig 1

GB 2 387 676 A

Fig 1

Fig 2

IMPROVED TV PROGRAMME SELECTION

This invention is directed to methods and systems for the assessment of television viewers' preferences for certain programmes and the recommendation of television programmes to individual viewers.

There are various known systems which purport to be able to recommend to a viewer which television programme from an available list they are most likely to want to watch. In order for such a system to function, some means of predicting the viewer's tastes is required, along with a system for recommending programmes based on those predicted tastes. Typically, some form of classification of the television programmes is employed, so that the preferences and recommendations are based on categories or types of programme.

Known systems are typically based on collaborative filtering, in which a population of viewers is assessed in conjunction each with each other. Typically this type of assessment involves recommending programmes to viewers whose viewing patterns fit a certain stereotype; for example, a habitual viewer of financial programmes may also be recommended golf programmes, if in the overall population there is a correlation between the watching of financial programmes and the watching of golf programmes. More generally, collaborative filtering arrangements recommend programmes to a viewer according to those programmes which have been chosen by other viewers with similar tastes; for example, a viewer who watches programmes A, B and C will be recommended programme D, which was watched by a significant number of other viewers who also saw A, B and C.

It is customary to improve the efficiency of collaborative filtering by entering for each viewer a number of parameters (such as age and gender) which have been shown to influence programme preference. Other systems rely upon direct viewer intervention through entry of programme ratings.

It has been proposed to improve the efficiency of collaborative filtering by the use of Bayesian nets and reference is directed in this regard to US patent 5,704,017.

It is found that viewer preferences are highly individualistic, evidencing subtleties that cannot (or can only with great difficulty) be distinguished through collaborative filtering. It would also be preferable to have a system that placed less rather than more reliance upon the direct intervention of viewers to establish personal preferences.

It is therefore an object of one aspect of the present invention to provide an improved system for recommending TV programmes to viewers which places no reliance upon collaborative filtering and which places reduced or no reliance upon direct intervention of viewers.

Accordingly, the invention consists in one aspect in a method of recommending television programmes to an individual television viewer, comprising the steps of: identifying at least one programme attribute; establishing for the programme attribute a plurality of programme classes; deriving membership functions for programmes comprising membership values indicating the degree of membership of, where appropriate, a plurality of classes; monitoring which of the programmes are watched by the individual viewer; determining from the membership functions of those programmes watched a preference profile for the individual viewer for the attribute as a function of the classes; comparing current and future programmes to the preference profile; and recommending to the viewer those available programmes whose membership functions most closely match the preference profile.

This allows the advantage that it is the preferences of the individual viewer which are being assessed, and upon which recommendations are based, rather than the preferences of a group of viewers sharing arbitrarily "similar" interests. This individually tailored technique in turn permits a far more accurate prediction of the viewer's tastes, and therefore produces better recommendations.

Advantageously, a plurality of programme attributes are identified and a preference profile determined for each viewer for each attribute; the step of comparing current and future programmes to the preference profile comprising the steps of comparing for each attribute the membership function of that programme with the preference profile of the viewer to derive a likelihood of

viewer preference for each attribute and the step of recommending comprising the steps of combining the likelihoods associated with the respective attributes. Suitably, the attributes include at least one nominal attribute and at least one ordinal attribute. Preferably, the classes of at least one nominal attribute are hierarchical.

In certain embodiments, the method comprises monitoring which programmes are rejected by the individual viewer, and modifying the preference profile in accordance with the membership functions of those rejected programmes.

Thus, it is not only those programmes which are chosen by the viewer, but also those which are rejected which are used to develop the picture of the user's tastes.

In another aspect, the invention provides a method of recommending television programmes to an individual television viewer, comprising the steps of: monitoring which programmes are selected for viewing by the viewer; determining for each selected programme that set of programmes which were available for viewing and were rejected in favour of the selected programme; determining from the selected programmes and the rejected programmes a preference profile for the viewer; comparing future programmes to the preference profile; and recommending to the viewer those available programmes which most closely match the preference profile.

Advantageously, the step of determining the preference profile comprises weighting the preference profile in respect of an attribute in favour of a selected programme or against a rejected programme in accordance with the extent to which the programme is represented by that attribute.

Suitably, the method comprises weighting the preference profile against a rejected programme to a lesser extent if the rejected programme is a repeat. In an embodiment, the method comprises weighting the preference profile against the attributes of a rejected programme to a lesser extent if the rejected programme is represented by a further attribute having a low preference for the viewer.

In yet another aspect, the invention consists in a method of recommending television programmes to an individual television viewer according to their individual preferences, comprising the steps of: classifying television programmes into a plurality of categories, which classification categories being common to a given population of viewers; monitoring which of the programmes are watched by the individual viewer; determining from the classifications of those programmes watched a preference profile for the individual viewer as a function of the classification categories; comparing current or future programmes to the individual preference profile; and recommending to the individual viewer those available programmes whose classifications most closely match the preference profile.

In still another aspect, the invention consists in a method of facilitating the recommendation of television programmes to an individual television viewer, comprising the steps of: monitoring which programmes are watched by the individual viewer to provide a preference profile; determining an available set of programmes; inputting said profile and said available set into a Bayesian network as respective nodes, and using the network to calculate the probability, $P(x)$, that a given programme of the available set would be chosen by the viewer, wherein the probability calculations performed between nodes in the network in order to determine $P(x)$ are weighted by values derived from axioms governing the relationships between nodes, which axioms having been determined independently of any viewer profile.

Preferably, the Bayesian network comprises a plurality of Bayesian nets, each corresponding to a respective attribute.

Advantageously, the step of monitoring comprises: determining for each programme watched that set of programmes which were available for viewing and were rejected in favour of the programme watched; and weighting the preference profile for that attribute to the extent that the rejected programme is represented by that attribute.

In certain embodiments, the method comprises weighting against the attributes of rejected programmes in proportion to the merit ascribed to those

attributes in the current preference profile. Suitably the method comprises weighting to a lesser extent against the attributes of a rejected programme if there exist reasons not to watch. In one embodiment, a reason not to watch is that the rejected programme is repeat. In another, a reason not to watch is that the rejected programme is represented by a further attribute having a low preference for the viewer.

Preferably, the said axioms take the form of probability equations.

In certain embodiments, at least one attribute is nominal and at least one attribute is ordinal. Suitably, the preference profile increases, decreases or is symmetrical with respect to changes in the value of an ordinal attribute, depending upon the nature of the attribute.

Advantageously, the said probability calculations are further weighted according to viewer input ratings of given programmes.

In one embodiment, the determination of the preference profile further comprises: at an initial start-up point providing to the viewer a list of programmes; receiving viewer input as to which programme of the list would be watched were the listed programmes available; and repeating these steps with different lists to determine an initial preference profile.

The invention will now be described by way of example with reference to the accompanying drawings, in which:

Figure 1 is a schematic diagram illustrating a television programme recommendation system according to an embodiment of the invention; and

Figure 2 is a diagram illustrating a Bayesian network employed in an embodiment of the invention.

An overview of the system according to embodiments of the invention is shown in Figure 1. It contains three main components:

1.   Programme Classifier: uses information describing currently available programmes, typically in the form of metadata description tags, to determine which classes a programme belongs to. The classifier contains a fuzzy classification algorithm, the function of which is described later.

2.   Viewer Profiler: actively monitors a viewer's TV viewing pattern in real-time and determines their viewing preferences from their viewing choices. It uses the Bayesian networks outlined below to adaptively learn these preferences.

3.   Programme Recommender: compares the current available set of programmes on the electronic programme guide (EPG) with the viewer's preferences and recommends those programmes that best match what the viewer is likely to want to watch.

The principal use for the invention is providing a recommendation of one television programme from an available set to a television viewer. The various parts of the system employs a variety of techniques, and in particular seek to:

- Characterise programmes using a set of orthogonal "attributes" that allow complex or ambiguous membership of classification categories;
- Learn viewer preferences for each programme attribute;
- Take account of availability of programmes;
- Take account of covariance and exchangeability amongst programme attribute sub-classes;
- Convert programme membership weights into watch probabilities and use them to estimate preference probabilities;
- Adaptively update preference probabilities for programme attributes;
- Learn viewer preferences over sets of orthogonal programme attributes;

- Learn viewer preferences from viewer supplied ratings of a programme viewed;
- Make ranked recommendations from a list of available programmes based on preferences;
- Make recommendations for ambiguously specified programmes.

These techniques, and other functions of the invention are described in more detail below. Though the system described herein is principally directed to the use of a single viewer, it should be noted that the invention is not restricted to such application. For example, in a family household, the system may be configured to recognize different family members, and recommend programmes according to their profiles. The options for such a variation will also be described below.

In the embodiments described below, it is assumed that all inferences are being done in real time for a single time period, t, for the set of programmes available for viewing at that point in time. These programmes can be thought of as competing for the viewer's attention. We also call these discrete "viewing events" (despite the fact that the user might not view anything). Of course, the invention may also be employed in non-real time situations.

The recommendation technique is based upon a system of programme characterisation or classification. Any given programme may be described by a set of "attributes", for instance, by genre, origin, language and violence. Each attribute has a number of 'states'. For example, the "genre" attribute may have states such as "comedy", "thriller" and "war", whilst the violence attribute has states such as "low", "medium" and "high". Thus, in order to classify programmes, any given programme may be described by its states for the various attributes. For example, a given programme may be described as:

Genre = "Comedy"

Origin = "Hollywood"

Language = "English"

Violence = "Low"

In the following, therefore, programmes are described as having states of some attribute, rather than in being assigned to some class or category. Some programmes may be characterised as belonging to more than one state for a given attribute. For example a programme may belong to Genre = "Romance" and Genre = "Comedy".

A programme is described by a family of "membership functions" that measure the degree of membership of that programme to each state for that attribute. For a given programme we might specify the membership functions as:

w(Genre = Comedy) = 0.8, w(Genre = War) = 0.1, w(Genre = Thriller) = 0.1

w(Origin = Hollywood) = 0.7, w(Origin = UK) = 0.3

w(Language = English) = 1.0, w(Language = Spanish) = 0.0

Evidently, a programme may have as many membership functions as there are programme attributes.

The system aims to "learn" a viewer's preference ordering for a particular attribute (such as "Genre") and assumes that, if the attributes are orthogonal, the preference orderings can be learned for each one separately. For each attribute, the viewer's preferences are assessed, and an attribute profile is determined. These attribute profiles are combined to form an overall profile of the viewer's preferences; a preference profile.

For a given attribute, a viewer would have a set of coefficients or probabilities-of-liking for each state $i, \ldots, n$, corresponding to their appreciation of these states as determined from their viewing habits. Thus for the attribute "Genre", a viewer might have a profile of $p(G_1) = 0.2$, $p(G_2) = 0.4$, and so on (where $1, 2, \ldots n$ are the ith states).

For a given attribute, programmes may be characterised at a higher level of ambiguity or abstraction than are allowed for by the set of states for that attribute. For example, in the *genre* attribute "comedy" may be an abstraction for a *number* of states, such as "black comedy", "light comedy", "slapstick comedy" etc. In fact full hierarchical classification may be employed, with as much detail in sub-classifications as is required to produce an accurate representation of a viewer's preferences.

It should be noted here that the system of hierarchical classification is such that if the evidence for a particular programme is entered at a state level, this may be translated into an attribute-level classification. Each of the states or sub-states in the hierarchy has an appropriate weighting towards its "parent" in the hierarchy. Each level in the hierarchy is typically normalized, though the value of each need not be equal. For example, if there were many different types of comedy, but very few different types of thriller, the comedy state would end up with a higher weighting. The end result is that no-matter how specific the state specified for a programme, a value for the top-level attribute may be determined.

In this embodiment, the attributes are assumed to be nominally scaled; ordinally scaled attributes are discussed later.

The preferences of a viewer are determined conditionally upon programmes being:

> watched;
>
> rejected/ignored;
>
> available to be watched.

Generally speaking, programmes that have been chosen and watched help reveal the preferences of the viewer. Programmes that have been considered for viewing and rejected also reveal preference, in a negative sense. However, the extent to which either of these events can help to determine preference will depend on whether the programme type is available for viewing in the first place, whether it is a repeat or not and the extent to which other attributes of the programme make it attractive.

In embodiments of the invention, the problem of assessing the viewer's preferences is modelled using Bayesian networks. The networks have a number of nodes representing factors affecting the likelihood of a programme being watched, and the nodes are linked by conditional probability relationships between them. For example, the probability that a programme will be watched is considered conditional on the attributes of that programme, and the viewer's preferences for those attributes. Thus the "programme watched" node might be expected to be linked to an "attribute" node and a "preference" node.

In the embodiment described below, a particular set of nodes and links is employed, though it should be recognized that the invention is not restricted to consideration of these particular factors, and the probabilistic relationships between them.

Figure 2 illustrates the network used in a particular embodiment. For each programme attribute, $k = 1, \ldots, m$, a preference node C, is used, with "state values" equal to the states of the attribute. For example, for the attribute 'genre' the state values of node C might be comedy, thriller, etc. All other nodes in the network are Boolean, having values true or false. They are explained as follows (where we have one of these for each of the states i of C):

W_i: this corresponds to a viewing event "watched" — this is true or false depending on whether the programme watched is characterised as being of state i for the given attribute.

M_i: "available" – this is true or false depending on whether a programme having state i is available or not.

E_i: "evidence" – this is true or false depending on whether there is evidence to support a preference for state i.

N_i: "reasons not to watch" – this is true or false depending on whether there are reasons not to watch something with state i.

R_i: "repeat" — this is true or false depending on whether the programme with state i is a repeat or not.

A_i: "other reasons not to watch" this is true or false depending on whether there are reasons (other than repeat) not to watch a programme with state i which are more concerned with states other than i.

Typically, at a "start-up" position of the system, the initial prior probability distribution for the hypothesis node, C (at time period t=0) is uniform, since when the system begins learning it knows nothing of the viewer's preferences.

The "reason not to watch" node (N_i) is conditional upon both the repeat (R_i) node and the external attribute (A_i) node. A viewer may simply have chosen not to watch a programme because it had been repeated on TV and he had seen it previously, thus the repeat nodes help us explain away why a programme was not watched. A large proportion of TV programmes are repeats, and therefore we should be careful not to give more emphasis to what rejection of a repeat can tell us about viewer preference over what rejection of a new programme can tell us.

The external attribute node reflects the fact that in updating the viewer's preference profile, there are various pitfalls. For example, when the viewer chooses not to watch a programme he may chose not to watch it because one or more of the attributes are unattractive. The problem with updating the beliefs about each attribute is that the system is typically unable to identify which attribute or combination of attributes was unattractive enough to cause the viewer to reject the programme. Of course this problem is an admission that the attributes may not be strictly orthogonal and that there may be conditional dependency on the others. We cannot model this conditional dependency directly since it would result in a cyclic graph. We instead model the dependency in separate BNs, as currently specified, and include external preference nodes in each BNs, again as currently specified, but do not connect these common nodes to create a cycle. The single most compelling reason for not watching a programme is then isolated (i.e. that attribute of the given programme least preferred in the current preference profile) and used to help explain why the viewer did not watch it and update the attribute preference nodes accordingly. This effectively results in the attribute states being updated unevenly in

accordance with the known prior relative preference between attributes as determined from our prior beliefs about the relative like or dislike between different states.

This technique has a number of advantages over a simpler "naïve" Bayesian network, where the viewer's preferences are assessed according to the frequency of "watch" and "¬watch" (i.e. "not watch") events.

Firstly, the treatment of fuzzy membership is more subtle, and hence more effective. A simple watched/not watched distinction does not account for the subtleties in the classification of programmes permitted here. When we reject a programme that has 100% membership of some state then we can simply set $p(W_i = \neg\text{watch}) = 1$. On the other hand a programme with 0% membership of some state is treated as "unavailable" to be watched. An unavailable programme state is equivalent to entering no likelihood change to the watch node because we have neither accepted nor rejected it. So $p(W_i \mid M_i = \text{unavailable}) = p(W_i)$.

We can therefore take a fuzzy membership function and transform it into a series of membership vectors each of which takes the state for $M_i = \{\text{available}, \text{unavailable}\}$. For example the fuzzy membership function for:

w(Genre = Comedy) = 0.7, w(Genre = War) = 0.2, w(Genre = Thriller) = 0.1

would be transformed into the following likelihood vectors for each state:

$$p(M_{comedy} = \text{available}) = w(\text{Genre} = \text{comedy}) = 0.7$$
$$p(M_{comedy} = \text{unavailable}) = 1 - w(\text{Genre} = \text{comedy}) = 0.3$$
$$p(M_{war} = \text{available}) = w(\text{Genre} = \text{war}) = 0.2$$
$$p(M_{war} = \text{unavailable}) = 1 - w(\text{Genre} = \text{war}) = 0.8$$
$$p(M_{thriller} = \text{available}) = w(\text{Genre} = \text{thriller}) = 0.1$$
$$p(M_{thriller} = \text{unavailable}) = 1 - w(\text{Genre} = \text{thriller}) = 0.9$$

If a programme is unavailable we can infer that we learn nothing from it — this is equivalent to applying a uniform likelihood distribution to W_i. If a programme is 100% available this results in the crisp update of W_i to either W_i = watched or W_i = not watched.

The evidence node, E_i, governs whether the profile is updated or not with the information currently in the BN; the updating of the profile is discussed later. If a programme is unavailable the profile is updated by a uniform distribution, regardless of whether we have watched it or not. This ensures that accepting or rejecting programmes results with zero availability on a state does not change the watch probability for that state. Likewise, rejecting a programme with partial availability (membership) results in a proportional decrease in our watch probability and a corresponding decrease in our preference for the state once evidence propagation has taken place. Watching a programme with partial availability (membership) results in an increase in evidence to support that state preference.

Another advantage of this system is its dealing with repeated rejection. The notion of availability is used to treat the rejection of a number of programmes of the same state as being equal to rejecting one member of that state. This ensures that the attributes programmes are not repeatedly downgraded, so as not to bias the preference profile against those attributes "unfairly". For example, someone who likes war programmes may not consider five programmes with 20% war content to be "equivalently" attractive and so would not expect "war" to be down rated by as much as when rejecting a single complete war film, $w(M_i)$ = 1, if they had chosen to reject five partial war films, each with $w(M_i)$ = 0.2.

In preferred embodiments, axioms are used to govern the relationships between nodes in the BNs. Such axioms are typically conditional probabilities or rules which allow the BN to function more efficiently. These rules are often obvious at a glance, but this obviousness is not apparent to the BN, which merely computes an answer from probabilities input to it. It has no notion, for example, that given available programmes A and B, if a user prefers A, he will likely watch A rather than B. Such logical notions must be enforced upon the BN.

These axioms are implemented throughout the BN. For example, in the nodes handling membership, axioms ensure that accepting or rejecting programmes with zero membership of a given state do not change the watch probability for that state.

The axioms are derived independently of any viewer preference profile. They are input to the BN by setting values for the conditional probability relationships between nodes. For example, the probability that a programme will be watched, given that it is preferred, and that there are no reasons not to watch it, may be set high, e.g. at 0.9. In contrast, if the programme is not preferred, the value may be set low.

In embodiments of the invention, two different categories of attribute, ordinal and nominal, are considered. An example of a nominal-type attribute is the "genre" attribute considered above, with states such as "war" and "comedy". With nominal types the states are different but with ordinal types states are ranked. We use different approaches when scoring a programme against viewer preferences depending on whether the state is nominal or ordinal. In the ordinal case the BN for learning viewer preferences is the same as that used in the nominal case except that $p(W_i | N_i, C)$ has different values to reflect the ordinal nature of the preference structure being modelled. Here, all ordinal valued attributes are defined on an $\{1, 2, 3, 4, 5\}$ scale.

An example of an ordinal attribute is the amount of bad language in a programme; some viewers are indifferent towards bad language, but for other viewers, an increase in bad language increases the likelihood that the programme will not be watched. This type of attribute is termed "monotonic decreasing preference". Other examples are "monotonic increasing preference", where some users are indifferent to an attribute, but others are increasingly drawn to higher values of the attribute, and "symmetric preference", where some viewers prefer programmes with a large amount of the attribute, and others prefer programmes with as little as possible. Here again, axioms are employed in order to input this functionality into the BN.

An important feature of the BN according to the embodiments described is the ability to model exchangeability and correlations between states or sub-states. If sub-states are similar then a viewer might choose to view one sub-state if another correlated sub-state is not available. Axioms governing such behaviour dictate, for example, that A is chosen when B is unavailable, and vice versa, and that programme C is chosen over A and B in equal measure. Furthermore, if the viewer watches many A's and B's are never available, attribute B will not be down rated by too much because it is never on TV. If, the first time a B becomes available whilst an A is not available, and the viewer watches it we will end up with very closely ranked set of preference probabilities for the A and B sub-states. Of course if the viewer chose not to watch the B then he is effectively saying it is not exchangeable with A's and so the system will down rate the B accordingly.

The viewer's preference profile is updated as an observation of $p(C|\underline{e})$, that is the probability of a preference given the evidence vector, $\underline{e}$. The evidence vector is derived from the E_i and M_i nodes, i.e. considering whether an attribute is available in the programme, and whether there is any evidence for that attribute available. As described above, the availability of evidence (E_i) is dependent upon whether the programme was watched (W_i), and the membership (M_i) of the programme. In this embodiment, $p(C|\underline{e})$ is calculated from the various nodes and conditional probability relationships in the BN shown in Figure 2, as a function of the factors listed therein, employing standard probability theory and equations, such as Bayes theorem. Of course, different calculations for updating the preference profile are possible.

The preference ordering BN according to the embodiments described preferably "learns" the viewer's preferences over a set of viewing events. It does so adaptively by replacing the prior distribution, at time t+1, of the hypothesis node with the posterior distribution calculated at time, t, where t is the measure of time in terms of viewing events.

In an embodiment of the invention, viewer ratings are received, in order to increase the accuracy and/or efficiency of the system. The easiest and most

straightforward way of assessing likes and dislikes is by asking the viewer to rate a programme immediately after the programme is finished. Once we have this information we can choose to update the preference ordering BNs with the appropriate membership weight functions for the programme. Of course, the simplest case for a viewer to express dislike is to change channel, if this is done we assume that the programme was rejected and do not require to solicit a rating; the viewer can of course enter one in should they wish to do so.

In general where a viewer likes a programme evidence is entered into the preference ordering BNs a number of times proportional to the degree of enjoyment. For instance a "good" viewer response might mean that the enjoyment of watching the programme was equivalent to casually watching the programme five times.

Different types of viewing may be assessed. We can define casual viewing as the act of watching a complete (i.e. say 90% of it) TV programme but where the viewer could not be bothered to rate it. Partial viewing would be where the viewer watched it but changed channels or turned the TV off. We may elicit like/dislike ratings when the viewer rates the programme at its end or during the programme (the viewer might not want to finish watching a "bad" programme). This will tend to be a conscious act either in response to a prompt or will be initiated by the viewer. The values input into the BNs may reflect these types, for example, by adding a coefficient to the rating.

Care must be taken over the updating of the preference profile in response to such ratings, as the preferred or disliked aspect may not constitute the entire membership function of the programme. For instance, in an embodiment, if a viewer disliked a programme with a given membership function, it is assumed that they would have liked to have watched the "opposite" programme, having the inverse membership function, the states being updated as appropriate to this opposite function.

Given the system for the assessment of a viewer's preferences, including BNs for the various attributes identified, an overall profile is achieved, which may be used to provide a recommendation to the viewer. This recommendation is

made from a calculation of the probability that the viewer would like to watch a particular programme. Since each attribute is treated as orthogonal the recommendation simply involves calculating the marginal probability of watching each programme, $p(W_j)$, using the multiplication rule.

In a preferred embodiment, the following steps are followed in recommending programmes and updating the preference profile:

1. Determine whether viewer-rating is positive or negative. Apply a viewer-rating algorithm to produce an orthogonal programme membership function if the viewer dislikes the programme. Determine from rating the number of times to repeat steps 2 – 10.

2. For the watched programme enter the membership function evidence into each membership node, M_i, thus: $p(M_i \mid e_{M_i}) = w(M_i)$ and $p(\neg M_i \mid e_{M_i}) = 1 - w(M_i)$

3. Using all rejected programmes available at time t (excluding the one watched) we calculate the membership posterior on the M_i nodes using the availability metric:

$$p(M_i = \text{available} \mid e_{M_i}) = 1 - \prod_{j=1}^{P}(1 - w(M_{ij}))$$

4. If watched programme contained programme or partial programme states that have been watched set $p(E_i \mid e_{E_i}) = \text{true}$ for these states.

5. If rejected programmes contained programme or partial programme states that have not been watched at all set $p(E_i \mid e_{E_i}) = \text{false}$ for these states.

6. Apply recommendation algorithm with $p(A_i)$ and $p(R_i)$ set to uniform distribution. Get $p(A_i)$ posteriors.

7.   Repeat steps 2 – 5 and use the $p(A_i)$ posteriors.

8.   For the programme watched we set the prior for $p(R_i = \text{true}) = 1$ or $p(R_i = \text{true}) = 0$ as appropriate.

9.   For all programmes rejected at time t, excluding the one watched, we calculate the repeats priors on the nodes using the repeats metric:

$$r_i = \frac{\sum_{j=1}^{m} s_j}{m}$$

where $s_j = 1$ if R_i is true and 0 otherwise

10.   Calculate p(C|$\underline{e}$)

11.   Set the prior for the next time period to $p_{t+1}(C) = p_t(C \mid \underline{e})$.

In embodiments, the starting position of the system varies. Evidently, at time t=0, the BN driven system has no evidence upon which to base recommendations. Initially, the user may be prompted to rate a number of sets of programmes generated from an EPG, or to say which would be watched if available. The system would then have some evidence upon which to base recommendations.

Initially the recommendations could be divided into strong and weak recommends. The system could be given a "bedding down" period of a number of weeks before it starts to make recommends at all. A pre-set criterion or data set could be used to form a basic set of evidence for the BN, which would eventually be removed (or diluted into negligibility) when the system has been running for a significant period of time.

In certain embodiments, the recommendation is made in different ways:

- Personal channel: Here the system compares the day's EPG with the viewer profile and configures a personal channel containing an optimal set of programmes.

- Recommendation list for what's on now: the system simply examines all programmes currently on (within ½ hour of each other) and lists those that it thinks the viewer might like to see.

- Interrupt warning: When watching a programme the system non-invasively interrupts the viewer to inform them of a "better" programme starting on another channel.

In a further embodiment, two types of programme availability are accounted for: "on-demand (OD)" programming and "always on (AO)" rolling channels.

OD programmes are available at all times, such as pay-per-view movies. The above approach might penalise against OD programmes because the viewer will repeatedly reject them. In this embodiment, we count OD items as being available when the viewer has browsed the OD section of the EPG and decided to watch/not watch a programme. However, these OD items are competing only against each other and the programmes that are on TV at the booked time of viewing. Therefore this restricted set may form the availability set for learning.

Furthermore, as OD programmes are normally repeats, in this embodiment a programme may be treated as a repeat only if the viewer has watched it (as opposed to treating it as a repeat if it has been available before). In order to provide such a function, an apparatus according to an embodiment of the invention employs a facility for "remembering" certain types of viewer selection, in order to aid future updating of the preference profile. Of course, such memory of selections is not limited to applications regarding OD programming or repeats, and indeed may be used advantageously with many other embodiments described herein.

AO rolling channels like News, Music and Shopping also run the risk of being continuously rejected in the described approach — after all if a viewer watches a film rather than a continuous news channel it doesn't mean he dislikes the News; he just may not want to see the news at that time. We instead need to treat such programming differently. If we separate original programme segments and the "generic" programming we can recommend original programming in the normal way. The generic programming is always available and it never needs to be recommended because the viewer can always choose to watch it whenever he likes.

In a yet further embodiment, the viewer may flag programmes that they definitely want to see recommended regardless of their viewing patterns. This deterministic approach would be entirely separate from the existing recommendation scheme, although the recommender could still learn the viewer profile from the programmes watched.

In a still further embodiment, the system may determine from the programme being watched which of a set of viewers is currently viewing. Thus, the system would only make recommendations relevant to, and update their profile.

In one embodiment, a system is employed to interpret directly from a text-based programme listing for input to the Programme Classifier (1). Such a system would typically require the full details of the classification system employed by the classifier, profiler and recommender. Alternatively, or in addition, the system operates with a set of metadata description tags accompanying the respective programmes.

It will be appreciated by those skilled in the art that the invention has been described by way of example only, and a wide variety of alternative approaches may be adopted.

The methods and systems here described can be implemented in software on a variety of host hardware, resident at the location of an individual viewer or – in appropriate applications – at a remote location. The software can usefully be incorporated or associated with electronic program suite software.

CLAIMS

1.   A method of recommending television programmes to an individual television viewer, comprising the steps of: identifying at least one programme attribute; establishing for the programme attribute a plurality of programme classes; deriving membership functions for programmes comprising membership values indicating the degree of membership of, where appropriate, a plurality of classes; monitoring which of the programmes are watched by the individual viewer; determining from the membership functions of those programmes watched a preference profile for the individual viewer for the attribute as a function of the classes; comparing current and future programmes to the preference profile; and recommending to the viewer those available programmes whose membership functions most closely match the preference profile.

2.   A method according to Claim 1, wherein a plurality of programme attributes are identified and a preference profile determined for each viewer for each attribute; the step of comparing current and future programmes to the preference profile comprising the steps of comparing for each attribute the membership function of that programme with the preference profile of the viewer to derive a likelihood of viewer preference for each attribute and the step of recommending comprising the steps of combining the likelihoods associated with the respective attributes.

3.   A method according to Claim 2, wherein the attributes include at least one nominal attribute and at least one ordinal attribute.

4.   A method according to Claim 3, wherein the classes of at least one nominal attribute are hierarchical.

5.   A method according to any one of the preceding claims, further comprising monitoring which programmes are rejected by the individual viewer, and modifying the preference profile in accordance with the membership functions of those rejected programmes.

6.   A method of recommending television programmes to an individual television viewer, comprising the steps of: monitoring which programmes are selected for viewing by the viewer; determining for each selected programme that set of programmes which were available for viewing and were rejected in favour of the selected programme; determining from the selected programmes and the rejected programmes a preference profile for the viewer; comparing future programmes to the preference profile; and recommending to the viewer those available programmes which most closely match the preference profile.

7.   A method according to Claim 6, wherein the step of determining the preference profile comprises weighting the preference profile in respect of an attribute in favour of a selected programme or against a rejected programme in accordance with the extent to which the programme is represented by that attribute.

8.   A method according to Claim 6 or Claim 7, comprising weighting the preference profile against a rejected programme to a lesser extent if the rejected programme is a repeat.

9.   A method according to any one of Claims 6 to 8, comprising weighting the preference profile against the attributes of a rejected programme to a lesser extent if the rejected programme is represented by a further attribute having a low preference for the viewer.

10.   A method of recommending television programmes to an individual television viewer according to their individual preferences, comprising the steps of: classifying television programmes into a plurality of categories, which classification categories being common to a given population of viewers; monitoring which of the programmes are watched by the individual viewer; determining from the classifications of those programmes watched a preference profile for the individual viewer as a function of the classification categories; comparing current or future programmes to the individual preference profile; and recommending to the individual viewer those available programmes whose classifications most closely match the preference profile.

11.   A method of facilitating the recommendation of television programmes to an individual television viewer, comprising the steps of: monitoring which programmes are watched by the individual viewer to provide a preference profile; determining an available set of programmes; inputting said profile and said available set into a Bayesian network as respective nodes, and using the network to calculate the probability, $P(x)$, that a given programme of the available set would be chosen by the viewer, wherein the probability calculations performed between nodes in the network in order to determine $P(x)$ are weighted by values derived from axioms governing the relationships between nodes, which axioms having been determined independently of any viewer profile.

12.   A method according to Claim 11, wherein the Bayesian network comprises a plurality of Bayesian nets, each corresponding to a respective attribute.

13.   A method according to Claim 12, wherein the step of monitoring comprises: determining for each programme watched that set of programmes which were available for viewing and were rejected in favour of the programme watched; and weighting the preference profile for that attribute to the extent that the rejected programme is represented by that attribute.

14.   A method according to Claim 13, comprising weighting against the attributes of rejected programmes in proportion to the merit ascribed to those attributes in the current preference profile.

15.   A method according to Claim 13 or Claim 14, comprising weighting to a lesser extent against the attributes of a rejected programme if there exist reasons not to watch.

16.   A method according to Claim 15 wherein a reason not to watch is that the rejected programme is repeat.

17.   A method according to Claim 15 wherein a reason not to watch is that the rejected programme is represented by a further attribute having a low preference for the viewer.

18.   A method according any of the Claims 12 to 17, wherein the axioms take the form of probability equations.

19.   A method according to any of the Claims 12 to 18, wherein at least one attribute is nominal and at least one attribute is ordinal.

20.   A method according to Claim 19, wherein the preference profile increases, decreases or is symmetrical with respect to changes in the value of an ordinal attribute, depending upon the nature of the attribute.

21.   A method according to any of the Claims 11 to 20, wherein the probability calculations are further weighted according to viewer input ratings of given programmes.

22.   A method according to any of the above claims, wherein the determination of the preference profile further comprises: at an initial start-up point providing to the viewer a list of programmes; receiving viewer input as to which programme of the list would be watched were the listed programmes available; and repeating these steps with different lists to determine an initial preference profile.

INVESTOR IN PEOPLE

Application No:	GB 0208607.2	**Examiner:**	Ben James
Claims searched:	1-10	**Date of search:**	6 January 2003

Patents Act 1977 : Search Report under Section 17

Documents considered to be relevant:

Category	Relevant to claims	Identity of document and passage or figure of particular relevance
X, E	1, 10, 22	WO 02/067578 A RANKIN PAUL JOHN et al. (Abstract)
X	1, 10	EP 1192802 A SCHAFFER JAMES D (Abstract)
X	1, 5, 6, 10	EP 1155571 A GUTTA SRINIVAS (Abstract)
X, E	1, 10	WO 02/42959 A SCHAFFER JAMES D (Abstract)
X	1,10	WO 02/11445 A GUTTA SRINIVAS et al. (Abstract)
A		US 5704017 A HECKERMAN DAVID E et al (Column 9 bottom 3 lines)

Categories:

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.

Field of Search:

Search of GB, EP, WO & US patent documents classified in the following areas of the UKC^V:

G4A

Worldwide search of patent documents classified in the following areas of the IPC^7:

G06F, H04N

The following online and other databases have been used in the preparation of this search report:

EPODOC, WPI, JAPIO